Domain specific vs Generic Network Visualization: an evaluation with metabolic networks

Romain Bourqui, Helen Purchase, Fabien Jourdan

Domain specific vs Generic Network Visualization: an evaluation with metabolic networks

Romain Bourqui
LaBRI,
Unité Mixte de Recherche CNRS (UMR 5800)
351, cours de la Libération F-33405 Talence cedex,
France
bourqui@labri.fr

Helen C. Purchase
School of Computing Science, Lilybank Gardens,
University of Glasgow, G12 8QQ,
United Kingdom
helen.purchase@gla.ac.uk

Fabien Jourdan
UMR1089 Xénobiotiques INRA-ENVT
180 chemin de Tournefeuille
St-Martin-du-Touch, Toulouse
France
fabien.jourdan@toulouse.inra.fr

Abstract

Metabolic networks have been drawn manually for many years, and over time have developed representational conventions that make them familiar to biologists. With increasing current biological discoveries, these networks need to be frequently updated and modified, and automatic visualization algorithms are thus becoming a necessity. Many existing automatic graph layout algorithms exist, and it is not known whether such generic algorithms are sufficiently useful for biologists, or whether algorithms that specifically consider the existing representational conventions are necessary. No prior task efficiency evaluation studies have been performed on biological network visualizations. This paper reports on an experiment comparing the task efficiency of biologically relevant motif-search tasks using three layouts, two of which were produced using existing generic graph layout algorithms (Force Directed, Hierarchical), and one which was specifically designed to take existing metabolic representation conventions into account (MetaViz). Despite the search task favouring the easy identification of node connectivity in the Force Directed layout, the results showed no efficiency difference between Force Directed and MetaViz. We conclude that embodying the representational conventions in an automatic algorithm is not an impediment to task efficiency, and that some minor improvements to MetaViz would enhance its usefulness for biologists even further.

Keywords: Metabolic networks, evaluation, graph drawing, bioinformatics

1 Introduction

1.1 Motivation

Providing a helpful visualization tool in biology often requires finding a balance between usability and user expectations in terms of representational conventions. As in other fields, (e.g. integrated circuits (VLSI)), biologists have over many years defined representational constraints for biological network drawings (Michal 1998). For instance, in Figure 1 (Michal 1998), Gerhard Michel (who is best known for his wall chart of biochemical pathways (Michal, 1993)) defined some appropriate representational constraints and manually drew this view of a metabolic network. Representations of metabolic networks can be used to find sets of connected biochemical reactions (motifs) (Lacroix, et al., 2006), to highlight quantitative values on nodes and edges (Paley & Karp 2006), or to follow metabolite fluxes. It is important to note that the representations like the one shown in Figure 1 were not designed specifically for any particular tasks.

Drawing these networks by hand has become impossible since automatic experiments and genome annotations currently generate networks containing hundreds of nodes and edges (Karp, et al., 2000). Biological network drawing algorithms have therefore been defined (Becker & Rojas 2001, Wegner & Kummer 2005, Bourqui, R., et al., 2007), in particular being designed to generate drawings in accordance with biologists' representational conventions.

Much work has also been done on the generation of visualizations of abstract networks within the graph drawing research community (Battista, et al., 1999, Kaufmann & Wagner 2001). The issue addressed in this paper is whether such existing generic algorithms should be recommended to biologists for the display of metabolic networks, or whether domain-specific layout algorithms that respect the representational conventions familiar to biologists should be used instead. This is an important question: if it is the case that generic

algorithms produce equivalent performance to domain-specific ones, this would indicate an advantage in developing methods that follow existing biologists' representational conventions. On the other hand, if generic algorithms produce better performance, biologists may abandon their commitment to these conventions in the interests of efficiency.

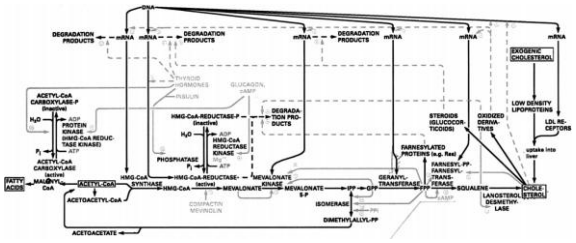

Fig. 1. Manual representation of a biochemical network (regulation mechanisms of cholesterol synthesis (Michal, 1998)).

This paper reports on an empirical study which compared the effectiveness of three layout algorithms when used with a motif-search task. Two of the layouts (Force Directed, Hierarchical) are popular existing generic layout algorithms; the third (MetaViz) is a layout specifically designed to include representational conventions familiar to biologists.

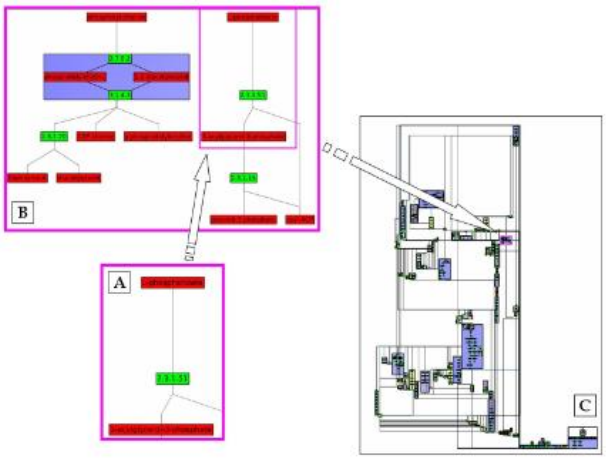

Fig. 2. The different abstractions of metabolic modelling. At the most detailed level, a metabolic reaction turns a metabolite (biochemical compound) into another one under the action of an enzyme (A). A set of metabolic reactions make up a metabolic pathway (B) which is a subgraph of the entire metabolic network (C).

1.2 Background

Our collaboration with biologists led us to focus on a particular biological research topic: metabolism. Metabolism is the set of biochemical reactions (figure 2.A) that are used to perform vital biological functions such as energy generation. Each metabolic function is modelled by a set of interconnected reactions corresponding to a small graph called a metabolic pathway (figure 2.B). Since the output of a pathway is often the input of another pathway it is possible to merge all these pathways into a single metabolic network (figure 2.C). Each organism has its own metabolic network. For instance, mammals and plants have different metabolic networks since only plants can generate energy using the photosynthesis pathway. Many metabolic networks exist and are updated regularly; automatic graph drawing algorithms for these pathways are therefore necessary.

Most of the work on metabolism visualization has been done at the pathway level (Becker & Rojas 2001, pp. 461–467, Schreiber 2003, pp. 105–110, Wegner & Kummer 2005). But in many metabolic studies it is necessary to visualize all the pathways and their connections at the same time (e.g. to put experimental data into context (Paley & Karp 2006)). Visualization is also necessary for topological analysis of metabolic networks, for example when looking for set of connected reactions (motifs) spanning over different pathways (Lacroix, et al., 2006, pp. 360–368). Simple pathway visualization is not sufficient for such tasks but neither is network visualization without pathway information. Indeed, to be useful for mapping experiments, it is necessary to represent the entire network structure while keeping the contextual information provided by its division into metabolic pathways. Note that this is one of the requirements for biological network visualization proposed by Saraiya et al. (2004). In the case of the motif search task, the drawing needs to provide a faithful image of the network structure. This is a challenging problem which we addressed by the development of the MetaViz layout method (Bourqui, R., et al., 2007).

MetaViz provides a domain specific solution for drawing the graph with its connected pathways. For our evaluation, we compared MetaViz with two generic layout methods.

To our knowledge, no prior work has been done on evaluation of biological networks layouts with respect to task efficiency. Saraiya et al. (2004) performed an informal heuristic evaluation of five popular pathway analysis systems, from which they identified requirements for pathway visualization systems. They did not, however, conduct a task-based experiment producing performance data.

2 Layouts

In this article, we present an empirical comparison of three different algorithms. First, we chose two classical graph drawing algorithms: a force directed algorithm and a hierarchical layout. Finally, we used our own algorithm which was specifically designed for metabolic network visualization (Bourqui R., et al., 2007).

2.1 Quotient graph modelling

Pathways are the building blocks of metabolic networks, and biologists need to visualize these features (Bourqui R., et al., 2007). Moreover some topological patterns like cycles are important since they correspond to particular biological processes (e.g. Krebs cycle for energy synthesis). Thus, a pre-processing step is defined before using any of the three layout algorithms; we applied a clustering algorithm (Bourqui, R., et al., 2007) to detect pathway and topological information. The result of this process is a quotient graph where nodes (metanodes) contain metabolic pathways or topological patterns. Two metanodes are linked by an edge (metaedge) if at least

two nodes (one in each metanode) are linked in the original network. The main disadvantage of quotient graph visualization is that it is not possible to know how many edges are represented by a given metaedge and which nodes within the metanodes are linked. Quotient graphs were used as the input to the three layout algorithms.

2.2 Force directed layout

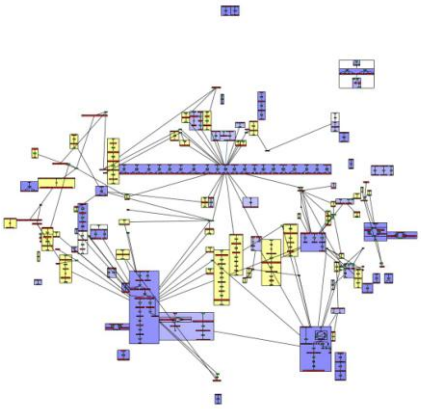

Fig. 3. Result of the force directed layout on the quotient graph.

Force directed layouts are widely used since they provide visually pleasing results which show the structure of the graph clearly. They behave as simulated physical systems which try to map the path distance between nodes in the network to euclidean distance and thus produce intuitive representations. There are several variations of this approach (e.g. Eades 1984, Frick, et al., 2004, Gajer & Kobourov 2000). We chose GEM (Frick, et al., 2004) since it gives particularly good results in term of stretch (i.e. the ratio between graph and euclidean distances) and is computationally efficient for the size of graphs we wished to use. To prevent node-node overlap, we first modified the algorithm by setting the ideal length of an edge to the sum of half the size of its extremities, and then used an algorithm (Dwyer, et al., 2005) to remove any remaining overlaps. Figure 3 shows an example of a quotient graph drawn using this method.

2.3 Hierarchical layout

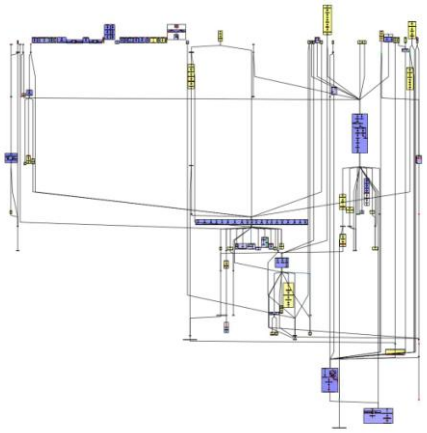

Fig. 4. Result of the hierarchical layout on the quotient graph.

The second type of algorithm we used is a hierarchical algorithm. This kind of algorithm embeds nodes on horizontal layers to highlight the hierarchical organization of data. This is followed by a heuristic which tries to minimize edge crossings by computing an ordering of the nodes on each layer. This type of algorithm is widely used in biological pathway drawings (Dogrusoz, et al., 2004, Karp, et al., 2002, Schreiber 2003). Like the force directed approach, many hierarchical algorithms exist (e.g. Sugiyama & Misue 1991, Auber 2003, Eiglsperger, et al., 2004); we chose the algorithm proposed by Auber (2003) which is an improvement of the well known Sugiyama algorithm (Sugiyama & Misue 1991). Figure 4 shows the result of this hierarchical algorithm on a quotient graph.

2.4 MetaViz layout

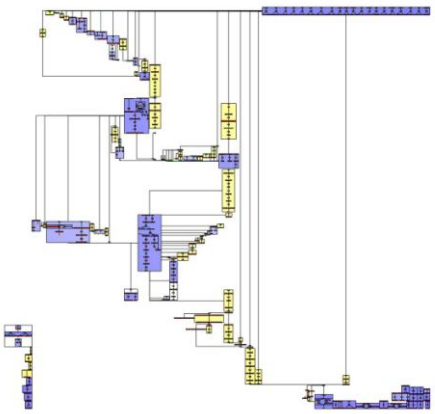

Fig. 5 Result of the MetaViz layout on the quotient graph

MetaViz (Bourqui, R., et al., 2007) is based on the Mixed Model algorithm of Gutwenger and Mutzel (1998). To adapt the Mixed Model algorithm to include metabolic network drawing conventions, we made three main modifications.

First, as we used the Mixed Model to draw the quotient graph, the algorithm was modified to take into account the varying sizes of metanodes.

Second, the Mixed Model is a planar graph drawing algorithm, so we needed to planarize the quotient graph. This problem is well-known and is NP-Hard (Lui & Geldmacher 1977). Many techniques exist, either by augmentation or by deletion of edges or nodes (Liebers 2001). The disadvantage of an augmentation based technique is that it may add up to $|V|^4$ nodes, with the drawing becoming difficult to understand. We therefore used the following heuristic: vertices of higher degree are removed one by one until the graph becomes planar. All removed nodes are then reinserted. Removed edges are re-added one by one as long as the graph is planar. The result of this process is then drawn by the modified Mixed Model algorithm. Finally, we add the edges removed during the planarization step. These edges are laid out on the external face of the drawing and with at most three bends per edge, in an orthogonal manner. This routing was inspired by hand-drawn representations of biological networks (e.g. Figure 1).

The third modification is related to the ordering of nodes. The Mixed Model algorithm has two steps:

1. The first step builds an ordered partition of the set of nodes. This partition is called shelling ordering. The principle used is that nodes that are on the external face of the graph are successively removed.
2. The second step is the recomposition of the graph according to the shelling ordering. To guarantee there is neither edge-edge crossing nor node-edge overlapping, the ordering is traversed in reverse order.

One of the metabolic network drawing conventions is that a reaction (or a compound) of a given metabolic pathway is embedded close to the other reactions (and compounds) of the pathway. The third modification of the Mixed Model therefore was the addition of a pathway constraint to the decomposition phase. If SO = {V1,V2, ...,Vr} is the shelling ordering, where each Vi is a set of nodes, when a vertex u is added into a set Vi , $1 \leq i < r$, we would like those nodes in the same pathways as u to be in Vi or Vi+1. However, the Mixed Model shelling ordering rules may prevent this. We therefore put these constrained nodes into the next possible Vj where $j > i$. Those nodes will then be more likely to be drawn next to each other. Figure 5 shows the result of the MetaViz layout on a quotient graph.

3 Methodology

3.1 Networks and tasks

We chose three different metabolic networks of different organisms. These networks are built with version 10.0 of the BioCyc database. Our collaborator, Ludovic Cottret, used perl scripts and pathway tools software (Karp, et al., 2000, Karp, et al., 2002, Krummenacker, et al., 2005) to obtain information on the reactions, compounds and metabolic pathways involved in the metabolism of three different genus of bacteria called Buchnera: Buchnera APS (graph A), Buchnera aphidicola BP (graph B) and Buchnera aphidicola SG (graph C). We chose organisms with similar size metabolic networks (503 nodes/526 edges, 558 nodes/538 edges, and 562 nodes/559 edges) and similar topologies so that the experimental tasks would not be of widely differing complexity (Bourqui et al. (2007) provide a more detailed description of the metabolic data).

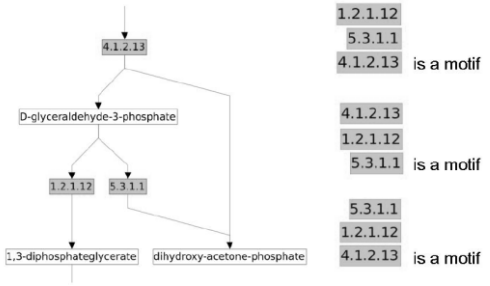

Fig. 6. Example of a motif where reactions are ordered in different ways. All three sets represent the same motifs. A motif is not necessarily a path.

Clustering (Bourqui, R., et al., 2007) was applied on each of these three networks to create the quotient graph, and three versions of each network were created, one for each of the three layout conditions. This resulted in nine graph drawings in total: these are referred to by their graph identifier (A, B, C) and their layout condition (GEM, Hierarchical, MetaViz).

The task is a biologically relevant one: the identification of motifs in networks. A motif is an unordered set of reactions such that each reaction of the motif shares (at least) one of its reactant or product with (at least) another reaction of the motif. Figure 6 shows an example of occurrence of the motif 1.2.1.12, 4.1.2.13, 5.3.1.1. Finding repeated motifs often indicate that gene duplications occurred during organism evolution.

Using the algorithm provided by Lacroix et al. (2006) we selected three motifs containing three reactions. To prevent our experimental participants being able to learn the answers, we chose motifs where the number of occurrences of the motif in the networks varied between 0 and 3 in the different organisms. We also selected motifs which could be found either within pathways or spanning over different pathways.

Table 1 shows the number of occurrences of each motif contained in each network (graph A, B and C), within a single metanode or shared by several metanodes.

Motifs	***#occurrences in graph A***		***#occurrences in graph B***		***#occurrences in graph C***	
	Within	Shared	Within	Shared	Within	Shared
6.3.4.* 3.5.4.9 6.3.2.17	0	2	0	2	0	0
4.3.2.* 6.3.4.5 2.1.3.3	1	1	1	1	1	1
2.7.2.* 1.2.1.13 2.2.1.1	0	1	0	1	0	0

Table 1: Number of occurrences of each motif contained by each network (graph A, B and C) within a single metanode or shared by several metanodes.

3.2 Experimental Design

Our evaluation used three layout algorithms, three different networks and three different motifs. Each task was therefore a combination of network, layout, and motif, with 27 tasks in total. The tasks were presented in random order.

Before commencing the experimental tasks, the participants completed 12 practice tasks chosen randomly from the 27 tasks. All participants preformed the same 12 practice tasks and therefore had the same experience at the beginning of the real experiment. During the first five practice tasks, the participants were helped by the experimenter and taught how to search for the relevant reactions. They were given feedback on their answers to these five tasks (Figure 7.(5)). For the following seven tasks, the participants were not aware that these were practice tasks and did not form part of the experimental data collection. The 27 experimental tasks were then presented in random order, and user-controlled rest breaks were included regularly throughout the duration of the experiment to address any problems of fatigue.

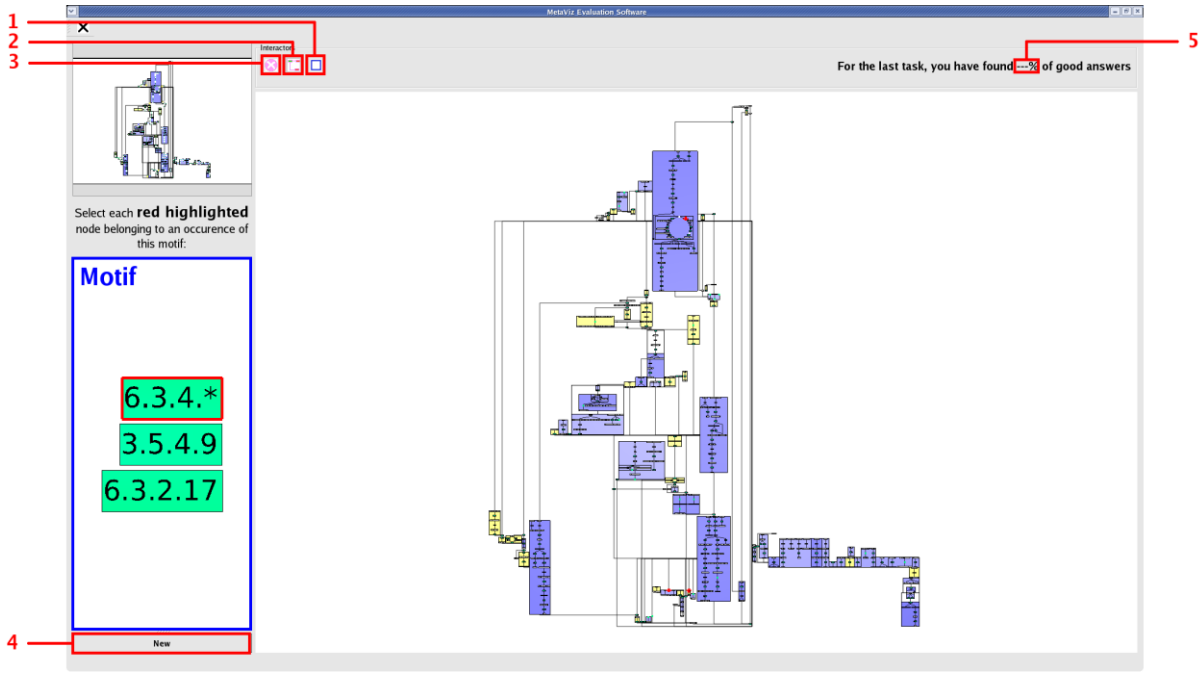

Fig. 7. Screenshot of the evaluation software. Buttons 1 and 2 allowed participants to select relevant reactions, with button 2 automatically highlighting the neighbourhood of the node in pink. Button 3 removed the pink highlighting. Button 4 was used to validate the answer and to move onto the next task. During the first five tasks, the participants had a feedback of their previous task, as a percentage of right answers (shown at 5).

3.3 Experimental task

Figure 7 shows a screenshot of the evaluation software. The visualization panel is located on the right, with the motif to search for on the left. To help the participants in their search, three hint nodes were highlighted in red in the network. These nodes were reactions potentially involved in the motif (e.g. all the nodes whose label starts with 6.3.4). The task consists in finding which of these hint nodes are part of (at least) one occurrence of the motif (here 6.3.4.*, 3.5.4.9, 6.3.2.17). The hints were necessary so as to prevent the user needing to search the whole of a very large network. Pilot tests revealed that the motif search task was still sufficiently challenging, despite the presence of these hints.

Using button 2 and clicking on a node automatically highlights in pink those nodes at distance of at most 2 from the selected node, and all edges and metaedges linking these nodes. Therefore to verify if a red highlighted hint reaction R is relevant, the participant had to click on it to see the reactions sharing at least one metabolite with R (Figure 8). It could be the case that only one other reaction R′ of the motif is found when looking at R's neighborhood. However, if R is the first reaction of a reaction cascade (a path), then the third reaction of the cascade would be at distance 4 from R. The participant would then need to look at the neighborhood of R′ to verify if R is relevant or not.

The participant then used the button 1 to select relevant reactions matching the motif (Figure 7). When the participants thought that they had found all the relevant nodes, the button 4 was used to validate this selection and to move on to the next task.

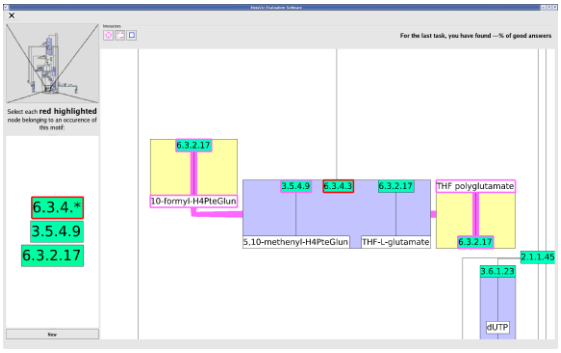

Fig. 8. Using button 2 and clicking on the reaction labelled 6.3.4.3 automatically highlights nodes at distance of at most two, and the edges and metaedges linking them. Here the reaction 6.3.4.3 is a relevant reaction since the other reactions of the motif (3.5.4.9 and 6.3.2.17) are highlighted in pink.

3.4 Experimental process

22 participants were recruited from Glasgow and Bordeaux Universities. Seven had some knowledge of bioinformatics; the others did not. The choice to exclude biologists from our sample was deliberate, and was motivated by an interview we had with 20 biologists. They were asked to order the three layouts according to their aesthetic expectations of metabolic network drawings. In 71% of the cases MetaViz was ranked first, in 29% it was ranked second, and it was never ranked third. Those participants who chose MetaViz indicated that it was the layout most familiar to them. Since we were interested in differences in task performance using these three layouts, independent of any prior familiarity, we deliberately did not include any biologists in our

sample, as we did not want to bias our results toward the MetaViz layout.

As this is a within-subject experiment, and participants' performance in one condition is compared with their own performance in another condition, any variation or similarity in the nature of participants does not affect the data analysis. The inclusion of practice tasks and task randomization helped counter any data bias due to the learning effect (whereby there is improved performance on the later tasks due to increasing task familiarity). Each experiment, including time spent at the beginning on the tutorial and the worked example, and on the questionnaire at the end, took approximately one hour. No problems were experienced during the experiments and all participants appeared to engage in the tasks seriously.

4 Results and analysis

The response time data for each task was measured as the time from the display of the network and the motif, to the time the participants pressed the "Validate" button to record that they had finished that task.

The error data was recorded as a 0 or 1, where 1 represents the case where the participant did not identify any of the present motifs correctly. Thus, a high value for both data measures (time and errors) implies poor performance. However, as the participants were allowed as much time as they wished to locate the motifs (or indicate an absence of motifs), there were very few errors in the responses. Hence, only the response time data is analyzed here.

4.1 Performance and results by Layout Condition

The average response time for the three layout conditions over all three networks and all three motifs is shown in Figure 9.

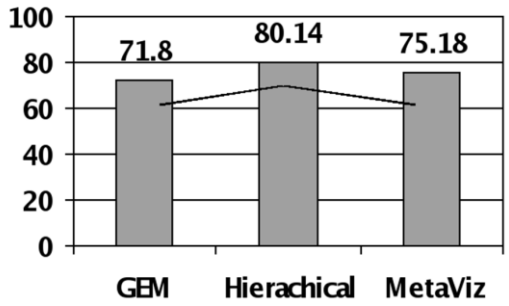

Fig. 9. The average response time in seconds for the three layout conditions, over all networks, and over all motifs. Lines indicate statistical significance between conditions at the 95% confidence level.

A two-tailed ANOVA test revealed statistical significance in performance over all conditions ($F=9.14>F(2,42,\alpha=0.05)=3.23$). Tukey pair-wise comparisons at the 95% confidence level led to the following conclusions:

1. The Hierarchical layout produces worst time performance than both the MetaViz and GEM layouts: an average of 80.14s (Hierarchical) versus 71.8s (GEM) and 75.18s (MetaViz).
2. There is no statistical difference in performance between the MetaViz and GEM layouts, despite the average for Metaviz (75.18s) being greater than that of GEM (71.8s).

There was no statistical difference in performance between the three networks A, B and C ($F=0.72<F(2,42, \alpha=0.05)=3.23$): this is as expected, as we chose networks of similar size and complexity. There was difference in the performance between the three motifs ($F=17.4>F(2,42,\alpha=0.05)=3.23$), with the first motif (6.3.4.*, 3.5.4.9, 6.3.2.17) being more difficult than both of the other two motifs. This is unsurprising, as this first motif included the most occurrences involving nodes shared between quotient nodes (see Table 1). No additional interesting results were obtained when the different layouts were compared within the data for each motif.

4.2 Preference results by Layout Condition

The post-task questionnaire asked the participants the following questions:

Q1. Which drawing is the best for the task?

Q2. Which drawing is worst for the task?

Q3. In which drawing is a highlighted edge easiest to follow?

Q4. In which drawing is a highlighted edge the most difficult to follow?

Q5. In which drawing is the neighbourhood of a node easiest to identify?

Q6. In which drawing is the neighbourhood of a node the most difficult to identify?

Participants were also invited to write textual comments on each of the three layouts.

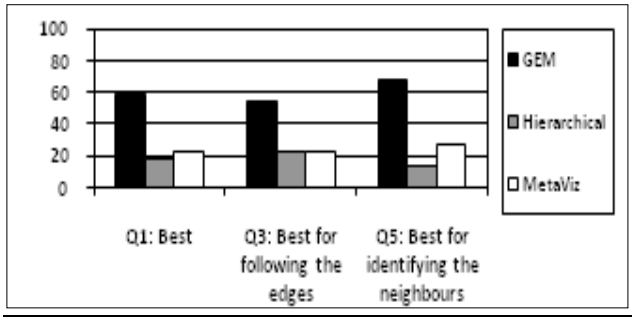

Fig. 10. Preference responses to the three post-task "best" questions, as percentages.

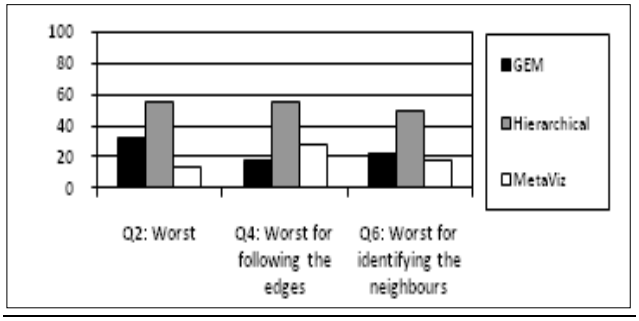

Fig. 11. Preference responses to the three post-test "worst" questions, as percentages.

Figures 10 and 11 show the percentage of the 22 participants who selected the layout conditions as best and worst, according to the six questions asked. Representative open comments from the participants regarding each of these layouts are shown in Table 2.

	Positive comments		Negative comments	
GEM	17	"short edges", "edges spaced out"	9	"node/edge overlaps"
Hierarchical	6	"no edge overlaps"	21	"long edges", "large graph area"
MetaViz	5	"orthogonal lines"	16	"long edges", "edges close together"

Table 2. Representative positive and negative comments about the three layouts.

We aggregated the quantitative preference data so that each participant effectively associated a score (between 1 and 3) to each of the three layouts. Using the time data for each participant for each layout, we performed a correlation analysis to see if there was any correspondence between preference and performance. There was no significance in this correlation data ($0.051<r(20, \alpha = 0.05)=0.4227$), indicating that the participants' preferred layout was not the one that they performed best on (and vice versa for their least preferred layout).

5 Discussion

Our expectation was that the generic Force Directed layout (Eades 1984, Frick, et al., 1994) would produce superior results in a motif-search task over both the other two algorithms because of the way in which it highlights connectivity.

Figure 12 shows three detailed views of the same metabolic network. Each view is obtained using one of the three algorithms, and is shown at the same zoom level. To highlight the connectivity of a node, we coloured all the paths of length two from that node. Using the same scaling factor (as in Figure 12), all the nodes at distance two from the focus node are visible under the GEM algorithm; this is not the case with the two other algorithms. With the Hierarchical and Metaviz algorithms, users would have to navigate the view (zoom in/out and pan) to view all the highlighted edges.

This example shows why we anticipated that for connectivity tasks GEM would provide better efficiency results. Force directed methods like GEM are designed to embed nodes that are close in terms of path length near to each other with respect to euclidean distance. In contrast, the Metaviz and Hierarchical layouts focus more on structuring the layout, node distributions, and avoiding edge crossings.

The data supports, to some extent, the hypothesis that GEM is superior, as the GEM layout results in better performance than the Hierarchical layout.

We were surprised, however, at the success of the MetaViz layout, whose performance was statistically as good as GEM. On looking at the MetaViz layout again, we believe this is because MetaViz as used with these networks has a clean appearance, with clear orthogonal lines and no edge or node overlaps. We also believe that the adaptation of the shelling ordering based on pathway constraints resulted in more compact node distributions, with higher information density in parts of the drawing. This is unlike the Hierarchical drawing, where the nodes are more dispersed. Thanks to our participant selection, we can affirm that the success of MetaViz cannot be attributed to prior biological knowledge or familiarity, as there were no biologists amongst our participants.

The preference data is the most telling when it comes to comparing the three layouts, as GEM is consistently rated the best (and never the worst), and Hierarchical is consistently rated the worst (and never the best). MetaViz is considered neither the best nor the worst.

There is an interesting anomaly in the reversal of the data between GEM and MetaViz for the overall "worst" question, Q2, where GEM is ranked the second worst (and therefore, by implication, the second best) by 31.8% versus 13.6%. Observation of the questionnaires showed that many of those participants who rated GEM the worst highlighted problems such as node/edge overlapping in their open comments.

The preference ranking order for the three layouts is therefore clearly GEM (best), MetaViz (middle), Hierarchical (worst). This contrasts with the performance data where GEM and MetaViz produce similar results.

We anticipated that GEM would be preferred because of the elegant layout aesthetics of the spring model and its depiction of close connections: this is supported by the textual remarks of the participants who commented favourably on the short edges and visual spread of nodes and edges.

However, our performance data shows that the MetaViz is just as effective as GEM, despite the fact that its layout does not appear to favour a connectivity task. Thus, as biologists typically prefer layouts similar to the MetaViz (which match the visualizations that they are familiar with), our data shows that doing so is not detrimental to their motif search efficiency when compared with the elegant GEM model favoured by researchers in graph layout.

6 Conclusions

When designing a metabolic network visualization tool the choice of the drawing algorithm is important since biologists expect particular representational conventions. Existing graph drawing algorithms like Force Directed and Hierarchical may prove useful in such tools. Our hypothesis was that on connectivity tasks users would be more efficient using Force Directed drawings. However our experimental results show that there is no efficiency difference between a diagram designed with biological conventions (Metaviz) and a Force Directed layout.

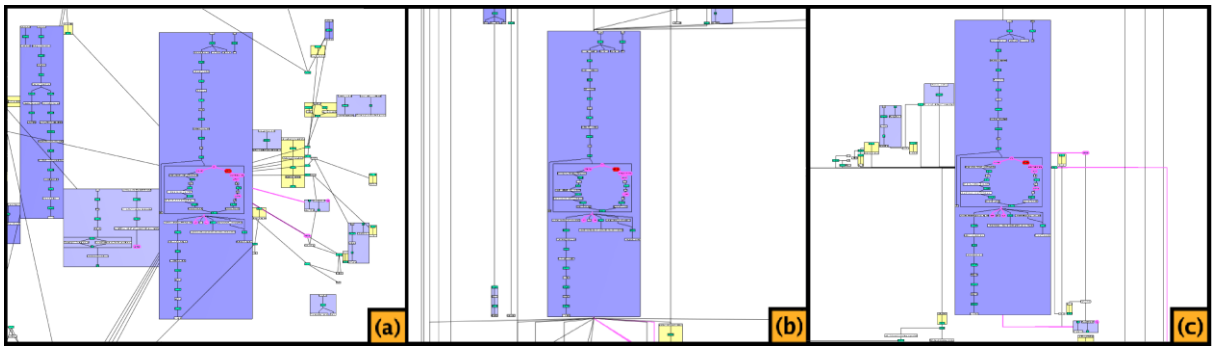

Fig. 12. Detail of a metabolic network drawn using GEM, Hierarchical and MetaViz algorithms. All the paths of length two going out of the red node are highlighted in purple.

We can conclude from these results that the efforts spent on layout algorithms that conform to biological representational conventions are worthwhile, because not only will such representations match biologists' expectations, they can be as efficient as generic spring-layout algorithms.

These results need, of course, to be interpreted within the context of this experiment and its limitations and parameters. The experiment used three networks of a particular size and three particular motifs. Using more than one network and more than one motif assists in producing generalizable results, but these are still constrained by the necessary limitations of the formal experimental method.

Using the formal experimental method allowed us to collect specific, measureable and controlled performance data associated with each of our three representations, thus enabling us to compare their effectiveness rigorously using statistical methods. While the experimental task we used may only be part of the activities typically performed on such visualisations, wider, more extensive exploration and communication tasks would not have been possible within this formal method. A more exploratory usability study could be envisaged which investigates these visualisations when used by experts with more extensive and richer real-world tasks: this would be an interesting further study. Such exploratory studies, however, do not produce clear and controlled data that can easily be analysed using statistical methods.

Since our aim was to evaluate the effect of drawing algorithms on user efficiency in a common biological motif searching task we chose non-biologist users: by removing the expectation of the participants having any domain knowledge, we could be sure that the task performance data truly represented the complexity of the visual motif-search task, and was not influenced by any prior biological knowledge. We anticipate that the same evaluation with biologists would confirm these results, and may show that MetaViz is superior since it includes representational conventions biologists would expect.

In addition, this experiment has provided us with useful qualitative data in the form of positive and negative comments about MetaViz. Integrating the suggestions made in our next version would ensure a much improved algorithm. For example, increasing the information density, removing white space and reducing the overall area of the diagram would address many negative comments received. These improvements will lead to even better experimental results.

7 Acknowledgements

We are grateful to: Ludovic Cottret for providing metabolic networks, the ANR-BBSRC funded project, Systryp, Glasgow Bioinformatics Research Centre and LaBRI, Bordeaux for hosting authors, and all the experimental participants. Ethical clearance for this study is covered by the student ethics protocol of the Department of Computing Science, University of Glasgow.

8 References

Auber, D. (2003): Tulip: A huge graph visualisation framework. In P. Mutzel & M. Jünger, eds., *Graph Drawing Softwares, Mathematics and Visualization*. Springer-Verlag, pp. 105-126.

Battista, G.D., Eades, P., Tamassia, R. & Tollis, I.G. (1999): *Graph Drawing Algorithms for the Visualization of Graphs*. Prentice Hall.

Becker, M. & Rojas, I. (2001): A Graph Layout Algorithm for Drawing Metabolic Pathways. *Bioinformatics*, **17**:461-467.

Bourqui R., et al., (2007): Metabolic network visualization eliminating node redundance and preserving metabolic pathways. *BMC Systems Biology*, **1(29)**.

Dogrusoz, U., Giral, E., Cetintas, A., Civril, A. & Demir, E. (2004): A compound graph layout algorithm for biological pathways. *Proc. Graph Drawing 2004 (GD'04)*, 442-447.

Dwyer, T., Marriott, K. & Stuckey, P.J. (2005): Fast node overlap removal. *Proc. Graph Drawing 2005 (GD'05)*, 153-164.

Eades, P. (1984): A heuristic for graph drawing. *Congressus Numerantium*, **42**:149-160.

Eiglsperger, M., Siebenhaller, M. & Kaufmann, M. (2004): An Efficient Implementation of Sugiyama's Algorithm for Layered Graph Drawing. *Proc. Graph Drawing 2004 (GD'04)*, 155-166.

Frick, A., Ludwig, A. & Mehldau, H. (1994): A Fast Adaptive Layout Algorithm for Undirected Graphs. *Proc. Graph Drawing 1994 (GD'94)*, 399-403.

Gajer, P. & Kobourov, S.G. (2000): GRIP: Graph dRawing with Intelligent Placement. *Proc. Graph Drawing (GD'00)*, 222-228.

Gutwenger, C. & Mutzel, P. (1998): Planar Polyline Drawings with Good Angular Resolution. *Proc. Graph Drawing 1998 (GD'98)*, 167-182.

Hinton, P.R. (1995): *Statistics Explained: A Guide for Social Science Students. Routledge*.

Karp, P., Paley, S. & Romero, P. (2002): The Pathway Tools software. *Bioinformatics*, **18(90001)**:S225–S232.

Karp, P., Riley, M., Saier M. & Paulsen, I (2000): The ecocyc and metacyc databases. *Nucleic Acids Research*, **28(1)**:56–59.

Kaufmann, M. & Wagner, D. (2001): *Drawing Graphs*. Springer.

Krummenacker, M., et al., (2005): Querying and computing with biocyc databases. *Bioinformatics*, **21(16)**:3454– 3455.

Lacroix, V., Fernandes, C.G. & Sagot, M-F. (2006): Motif search in graphs: application to metabolic networks. *IEEE/ACM Trans. Comput. Biol. Bioinform*., **3(4)**:360–368.

Liebers A. (2001): Planarizing Graphs - A Survey and Annotated Bibliography. *Journal of Graph Algorithms and Applications*, **5(1)**:1–74.

Lui, P. & Geldmacher, R. (1977): On the deletion of nonplanar edges of a graph. *Proc. on the 10th conf. on Comb., Graph Theory, and Comp*., 727-738.

Michal, G. (1993): *Biochemical Pathways (Poster).* BoehringerMannheim.

Michal, G. (1998): On representation of metabolic pathways. *BioSystems*, **47**:1-7.

Paley, S.M., Karp, P.D. (2006): The pathway tools cellular overview diagram and Omics Viewer. *Nucleic Acids Research*, **34**:3771–3778.

Saraiya, P., North, C. & Duca, K. (2004): Visualizing biological pathways: requirements analysis, systems evaluation and research agenda. *Information Visualization*, **4(3)**:1–15

Schreiber, F. (2003): Comparison of metabolic pathways using constraint graph drawing. *APBC 03: Proc. of the First Asia-Pacific bioinformatics conference on Bioinformatics*, 105-110.

Sugiyama K. & Misue, K. (1991): Visualisation of structural information : Automatic drawing of compound digraphs. *IEEE Transactions on Systems, Man, and Cybernetics*, **21(4)**:876–892

Wegner K. & Kummer U. (2005): A new dynamical layout algorithm for complex biochemical reaction networks. *BMC Bioinformatics*, **6(212)**.